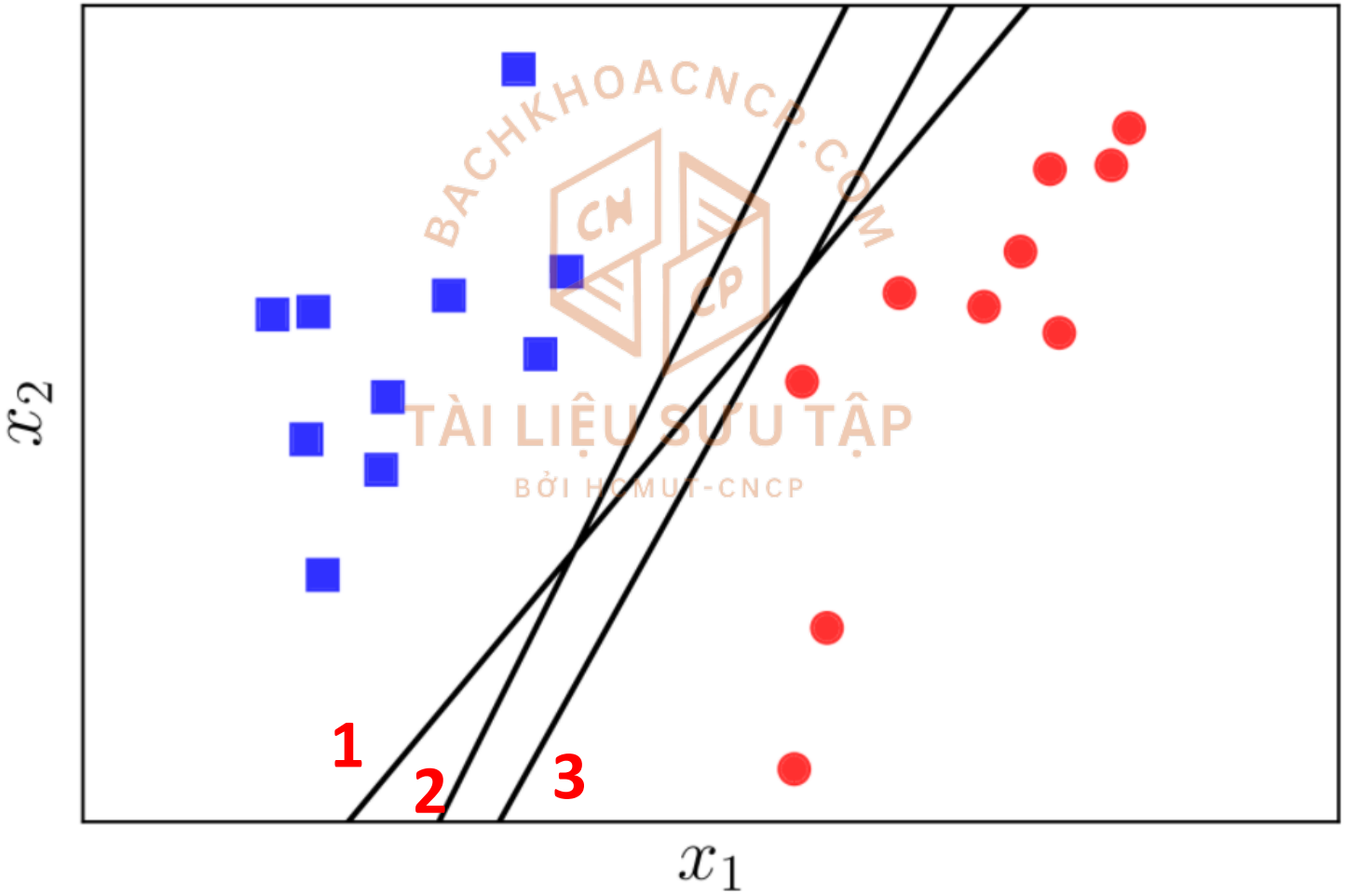
Artificial Intelligence

Soft margin Support Vector Machine

Pham Viet Cuong
Dept. Control Engineering & Automation, FEEE
Ho Chi Minh City University of Technology

Support Vector Machine

- ✓ Two-class classification problem with linearly separable data

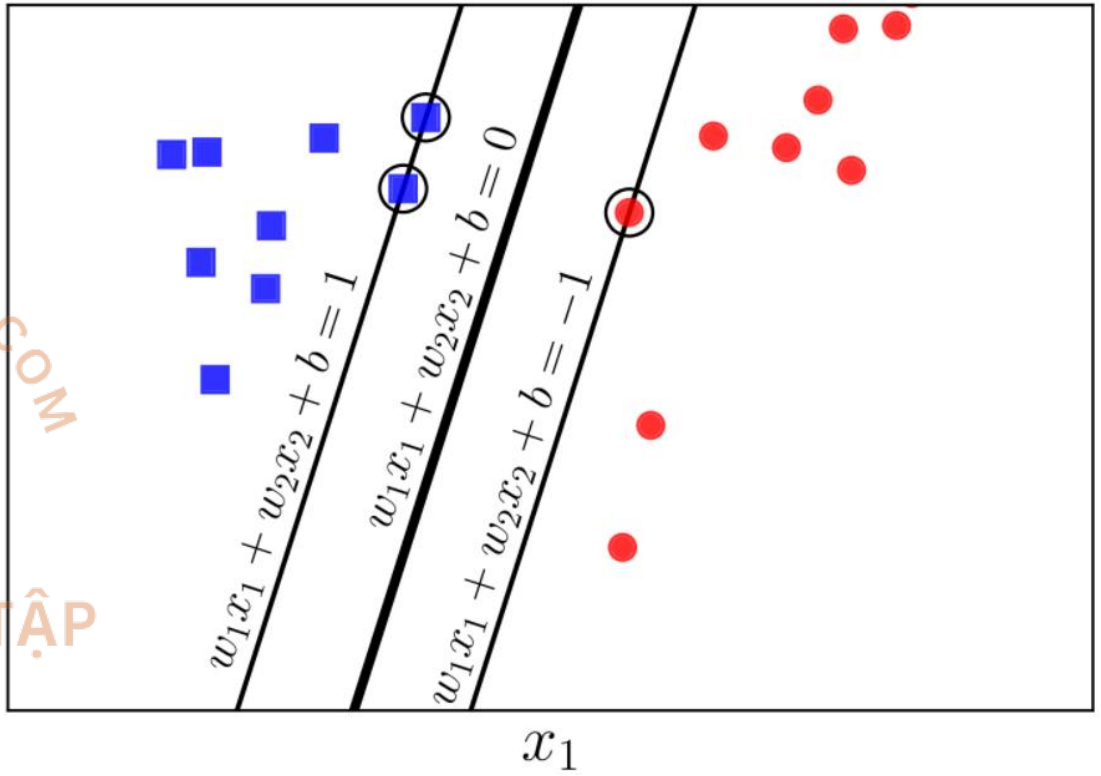


Support Vector Machine

- Optimization problem

$$(\mathbf{w}, b) = \arg\max_{\mathbf{w},b} \left\{ \min_n \frac{y_n(\mathbf{w}^T\mathbf{x}_n + b)}{||\mathbf{w}||_2} \right\}$$

$$\arg\max_{\mathbf{w},b} \left\{ \frac{1}{||\mathbf{w}||_2} \min_n y_n(\mathbf{w}^T\mathbf{x}_n + b) \right\}$$



- Assume that

$$\min \; y_n(\mathbf{w}^T\mathbf{x}_n + b) = 1$$

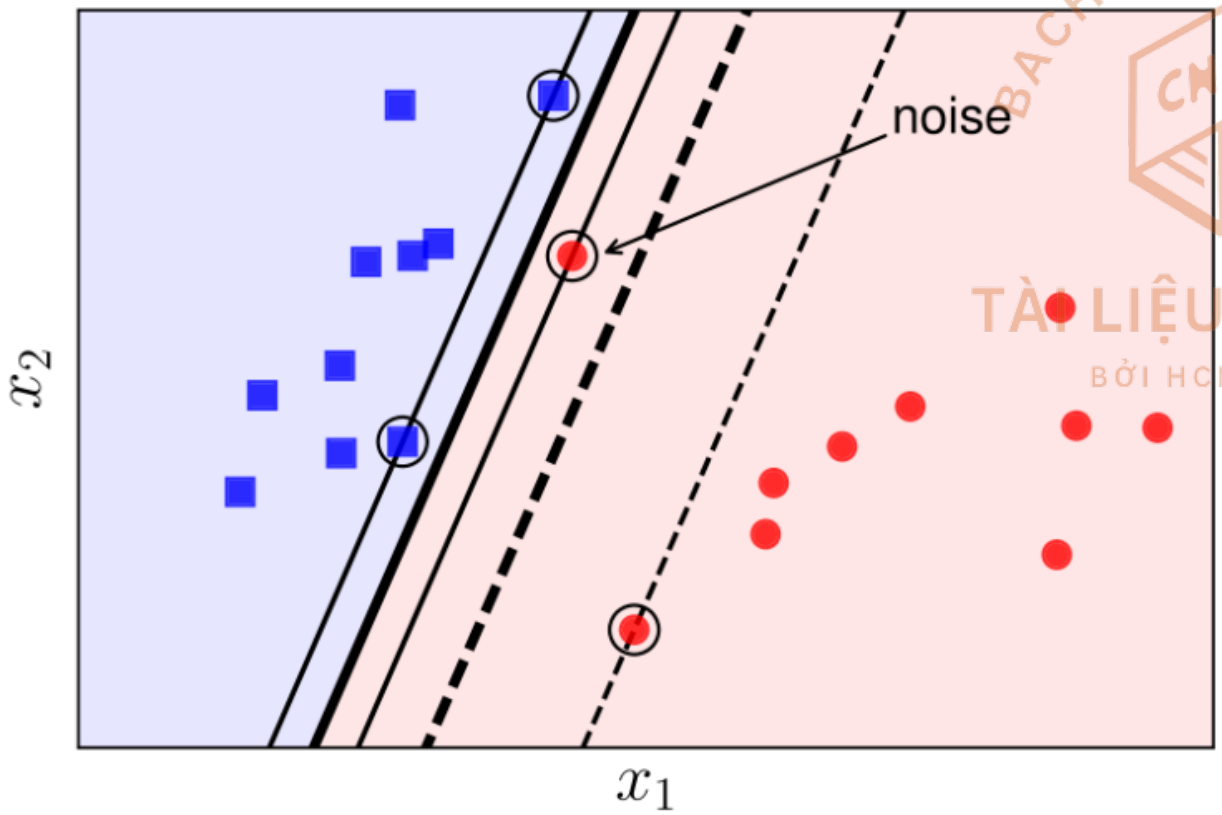
$$(\mathbf{w}, b) = \arg\max_{\mathbf{w},b} \frac{1}{||\mathbf{w}||_2} \text{ subject to: } \; y_n(\mathbf{w}^T\mathbf{x}_n + b) \geq 1, \forall n = 1, 2, \ldots, N$$

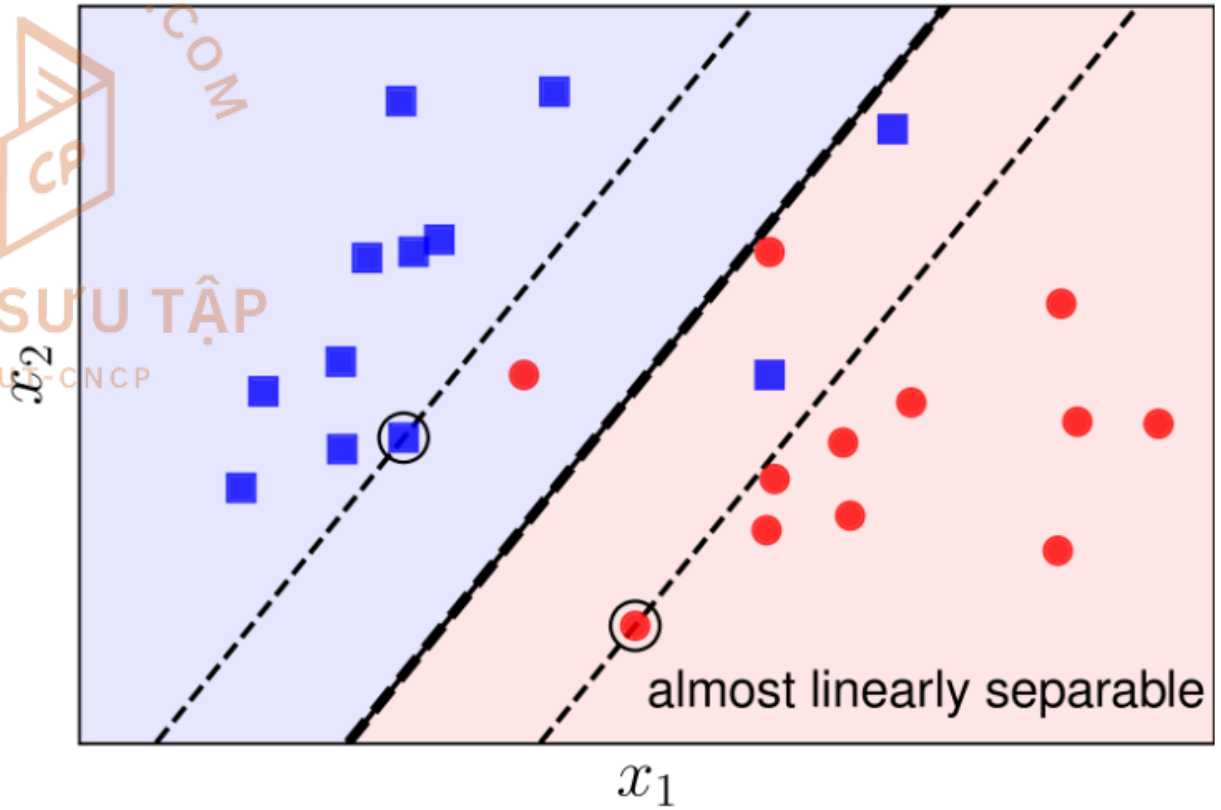
$$(\mathbf{w}, b) = \arg\min_{\mathbf{w},b} \frac{1}{2}||\mathbf{w}||_2^2 \text{ subject to: } \; 1 - y_n(\mathbf{w}^T\mathbf{x}_n + b) \leq 0, \forall n = 1, 2, \ldots, N$$

Soft margin SVM

- Two-class classification problem
  - Noise
  - Non linearly separable data

→ Soft margin SVM





Soft margin SVM

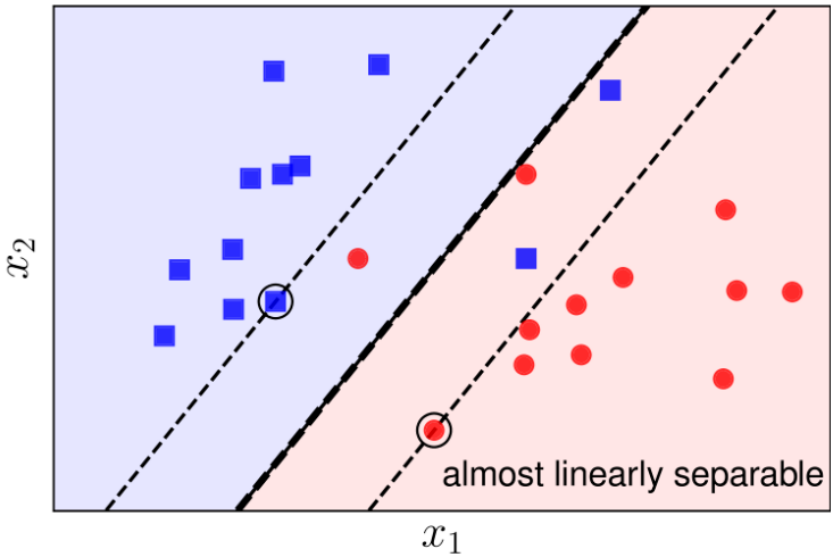
- Objective
  - Large margin
  - Small violations
- Slack variable
  - Without violation: $\xi_n = 0$
  - With violation:

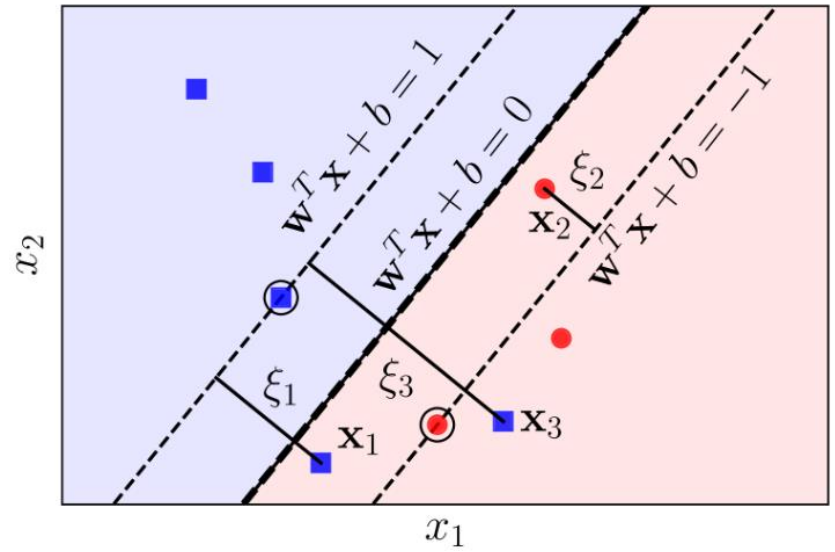
$$0 < \xi_n < 1$$

or

$$\xi_n > 1$$

$$\xi_i = |\mathbf{w}^T \mathbf{x}_i + b - y_i|$$





Soft margin SVM

- SVM $$(\mathbf{w}, b) = \arg\min_{\mathbf{w},b} \frac{1}{2}||\mathbf{w}||_2^2$$

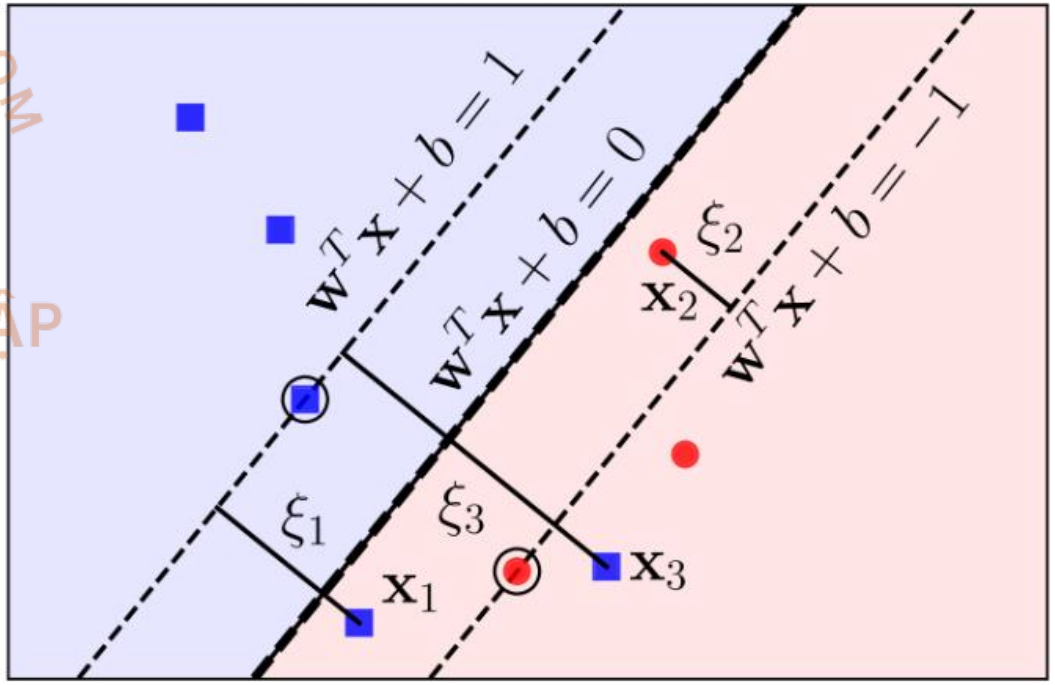
$$\text{subject to:} \quad y_n(\mathbf{w}^T\mathbf{x}_n + b) \geq 1, \forall n = 1, 2, \ldots, N$$

- Soft margin SVM
  - Objective functiontion:

$$\frac{1}{2}||\mathbf{w}||_2^2 + C\sum_{n=1}^{N}\xi_n$$

  - Constrain:

$$y_n(\mathbf{w}^T\mathbf{x}_n + b) \geq 1 - \xi_n$$



$$1 - \xi_n - y_n(\mathbf{w}^T\mathbf{x}_n + b) \leq 0, \quad \forall n = 1, 2, \ldots, n \qquad \xi_n \geq 0, \ \forall n = 1, 2, \ldots, N$$

Soft margin SVM

- Soft margin SVM
  - Objective functiontion

$$\frac{1}{2}||\mathbf{w}||_2^2 + C\sum_{n=1}^{N}\xi_n$$

  - How to choose C?